



# Technical Collaboration in the Web Archiving Community

WASAPI National Symposium  
February 22, 2017

How is the web archiving  
community doing in terms  
of collaboration?

# Current State

- Relatively big & hard problem
- Relatively small resourcing
- Lots of mindshare, passion and goodwill
- Relatively few shared solutions / collective approaches

# Current State

- Relatively big & hard problem
- Relatively small resourcing
- Lots of mindshare, passion and goodwill
- Relatively few shared solutions / collective approaches



# Prerequisites of Effective Collaboration

- Alignment of purpose, assumptions & resources
- Enlightened self-interest
  - People collaborate to help themselves
- Communication & collaboration tools
- Bite-sized tasks
  - A role for everyone



# 4 Steps towards a WA Community Now

## 1. Communication Channels

- Email List, Slack, Github

## 2. Face time

- We need an annual, North American web archiving meeting
- Web archivists, LAMS, technologists, researchers, data scientists & other users

## 3. Use Cases & Shared Architecture

- Shared understanding of objectives & methods

## 4. APIs

- Carve up the elephant: allow mix & matching
- Strategy for sustained definition and maintenance of common APIs

# The Power of APIs

- Componentize the problem: “carve up the elephant”
- Agree on standard interactions
- Interoperable, swappable components across
  - platform
  - institutions
  - place
  - time



# (Sketchy) Web Archiving APIs

- Capture
- QA
- Store
- Extract
- Export/Import
- Index
- Search
- **Preservation**
- **Playback**
- **Mining**
- **Monitoring**
- **Reporting**
- ...